

(19)



Europäisches Patentamt

European Patent Office

Office européen des brevets



(11)

EP 0 797 324 B1

(12)

EUROPEAN PATENT SPECIFICATION

(45) Date of publication and mention
of the grant of the patent:
24.11.2004 Bulletin 2004/48

(51) Int Cl.7: **H04H 5/00**

(21) Application number: **97301677.7**

(22) Date of filing: **12.03.1997**

(54) Enhanced joint stereo coding method using temporal envelope shaping

Verbessertes Kombinationsstereokodiervverfahren mit zeitlicher Hüllkurvenformgebung

Méthode améliorée de codage stéréo combiné utilisant une mise en forme d'enveloppe temporelle

(84) Designated Contracting States:
DE FR GB

(30) Priority: **22.03.1996 US 621762**

(43) Date of publication of application:
24.09.1997 Bulletin 1997/39

(73) Proprietor: **LUCENT TECHNOLOGIES INC.**
Murray Hill, New Jersey 07974-0636 (US)

(72) Inventor: **Herre, Juergen Heinrich**
Basking Ridge, New Jersey 07920 (US)

(74) Representative:
Buckley, Christopher Simon Thirsk et al
Lucent Technologies (UK) Ltd,
5 Mornington Road
Woodford Green, Essex IG8 0TU (GB)

(56) References cited:
EP-A- 0 563 832 WO-A-95/08227
WO-A-95/26083 US-A- 5 481 614

EP 0 797 324 B1

Note: Within nine months from the publication of the mention of the grant of the European patent, any person may give notice to the European Patent Office of opposition to the European patent granted. Notice of opposition shall be filed in a written reasoned statement. It shall not be deemed to have been filed until the opposition fee has been paid. (Art. 99(1) European Patent Convention).

Description

Cross-Reference to Related Application

[0001] The subject matter of this application is related to that of the U.S. Patent Application of J. Herre, entitled "Perceptual Noise Shaping in the Time Domain via LPC Prediction in the Frequency Domain," Ser. No. 08/585086, filed on January 16, 1996 and assigned to the assignee of the present invention and corresponding to EP-A-0 785 631 published on 23.07.97. "Perceptual Noise Shaping in the Time Domain via LPC Prediction in the Frequency Domain" is hereby incorporated by reference as if fully set forth herein.

Field of the Invention

[0002] The present invention relates to the field of audio signal coding and more specifically to an improved method and apparatus for performing joint stereo coding of multi-channel audio signals.

Background of the Invention

[0003] During the last several years so-called "perceptual audio coders" have been developed enabling the transmission and storage of high quality audio signals at bit rates of about 1/12 or less of the bit rate commonly used on a conventional Compact Disc medium (CD). Such coders exploit the irrelevancy contained in an audio signal due to the limitations of the human auditory system by coding the signal with only so much accuracy as is necessary to result in a perceptually indistinguishable reconstructed (*i.e.*, decoded) signal. Standards have been established under various standards organizations such as the International Standardization Organization's Moving Picture Experts Group (ISO/MPEG) MPEG1 and MPEG2 audio standards. Perceptual audio coders are described in detail, for example, in U.S. Patent No. 5,285,498 issued to James D. Johnston on Feb. 8, 1994 and in U.S. Patent No. 5,341,457 issued to Joseph L. Hall II and James D. Johnston on Aug. 23, 1994, each of which is assigned to the assignee of the present invention.

[0004] Generally, the structure of a perceptual audio coder for monophonic audio signals can be described as follows:

- The input samples are converted into a subsampled spectral representation using various types of filterbanks and transforms such as, for example, the well-known modified discrete cosine transform (MDCT), polyphase filterbanks or hybrid structures.
- Using a perceptual model, one or more time-dependent masking thresholds for the signal are estimated. These thresholds give the maximum coding error that can be introduced into the audio signal

while still maintaining perceptually unimpaired signal quality. In particular, these masking thresholds may be individually determined on a sub-band by sub-band basis. That is, each coder frequency band, which comprises a grouping of one or more spectral coefficients, will be advantageously coded together based on a correspondingly determined masking threshold.

- The spectral values are quantized and coded (on a coder frequency band basis) according to the precision corresponding to the masking threshold estimates. In this way, the quantization noise may be hidden (*i.e.*, masked) by the respective transmitted signal and is thereby not perceptible after decoding.

- Finally, all relevant information (*e.g.*, coded spectral values and additional side information) is packed into a bitstream and transmitted to the decoder.

Accordingly, the processing used in a corresponding decoder is reversed:

- The bitstream is decoded and parsed into coded spectral data and side information.
- The inverse quantization of the quantized spectral values is performed (on a frequency band basis corresponding to that used in the encoder).
- The spectral values are mapped back into a time domain representation using a synthesis filterbank.

[0005] Using such a generic coder structure it is possible to efficiently exploit the irrelevancy contained in each signal due to the limitations of the human auditory system. Specifically, the spectrum of the quantization noise can be shaped according to the shape of the signal's noise masking threshold. In this way, the noise which results from the coding process can be "hidden" under the coded signal and, thus, perceptually transparent quality can be achieved at high compression rates.

[0006] Perceptual coding techniques for monophonic signals have been successfully extended to the coding of two-channel or multichannel stereophonic signals. In particular, so-called "joint stereo" coding techniques have been introduced which perform joint signal processing on the input signals, rather than performing separate (*i.e.*, independent) coding processes for each input signal. (Note that as used herein, as used generally, and as is well known to those of ordinary skill in the art, the words "stereo" and "stereophonic" refer to the use of two or more individual audio channels.)

[0007] There are at least two advantages to the use of joint stereo coding techniques. First, the use of joint stereo coding methods provide for the ability to account for binaural psychoacoustic effects. And, second the required bit rate for the coding of stereophonic signals may

be reduced significantly below the bit rate required to perform separate and independent encodings for each channel.

[0008] Generally, the structure of a multi-channel stereophonic perceptual audio coder can be described as follows:

- The samples of each input signal are converted into a subsampled spectral representation using various types of filterbanks and transforms, such as, for example, the modified discrete cosine transform (MDCT), polyphase filterbanks or hybrid structures.
- Using a model, the time-dependent masking threshold of the signal is estimated for each channel. This gives the maximum coding error that can be introduced into the audio signal while still maintaining perceptually unimpaired signal quality.
- To perform joint stereo coding, portions of the spectral coefficient data are jointly processed to achieve a more efficient representation of the stereo signal. Depending on the joint stereo coding method employed, adjustments may be made to the masking thresholds as well.
- The spectral values are quantized and coded according to the precision corresponding to the masking threshold estimate(s). In this way, the quantization noise is hidden (*i.e.*, masked) by the respective transmitted signal and is thereby not perceptible after decoding.
- Finally, all relevant information (*i.e.*, the coded spectral values and additional side information) is packed into a bitstream and transmitted to the decoder.

Accordingly, the processing used in the encoder is reversed in the decoder:

- The bitstream is decoded and parsed into coded spectral data and side information.
- The inverse quantization of the quantized spectral values is carried out.
- The decoding process for the joint stereo processing is performed on the spectral values, thereby resulting in separate signals for each channel.
- The spectral values for each channel are mapped back into time domain representations using corresponding synthesis filterbanks.

[0009] Currently, the two most commonly used joint stereo coding techniques are known as "Mid/Side" (M/S) stereo coding and "intensity" stereo coding. The

structure and operation of a coder based on M/S stereo coding is described, for example, in U.S. Patent No. 5,285,498 (*see above*). Using this technique, binaural masking effects can be advantageously accounted for and in addition, a certain amount of signal-dependent gain may be achieved.

[0010] The intensity stereo method, however, provides a higher potential for bit saving. In particular, this method exploits the limitations of the human auditory system at high frequencies (*e.g.*, frequencies above 4 kHz), by transmitting only one set of spectral coefficients for all jointly coded channel signals, thereby achieving a significant savings in data rate. Coders based on the intensity stereo principle have been described in numerous references including European Patent Application 0 497 413 A by R. Veldhuis *et al.*, filed on January 24, 1992 and published on August 5, 1992, and (using different terminology) PCT patent application WO 92/12607 by M. Davis *et al.*, filed on January 8, 1992 and published on July 23, 1992.

[0011] By applying joint stereo processing to the spectral coefficients prior to quantization, additional savings in terms of the required bit rate can be achieved. For the case of intensity stereo coding, some of these savings derive from the fact that the human auditory system is known to be insensitive to phase information at high frequencies (*e.g.*, frequencies above 4 kHz). Due to the characteristics of human hair cells, signal envelopes are perceptually evaluated rather than the signal waveform itself. Thus, it is sufficient to code the envelope of these portions of a signal, rather than having to code its entire waveform. This may, for example, be accomplished by transmitting one common set of spectral coefficients (referred to herein as the "carrier signal") for all participating channels, rather than transmitting separate sets of coefficients for each channel. Then, in the decoder, the carrier signal is scaled independently for each signal channel to match its average envelope (or signal energy) for the respective coder block.

[0012] The following processing steps are typically performed for intensity stereo encoding/decoding on a coder frequency band basis:

- From the spectral coefficients of all participating channels, one "carrier" signal is generated that is suited to represent the individual channel signals. This is usually done by forming linear combinations of the partial signals.
- Scaling information is extracted from the original signals describing the envelope or energy content in the particular coder frequency band.
- Both the carrier signal and the scaling information are transmitted to the decoder.
- In the decoder, the spectral coefficients of the carrier signal are reconstructed. The spectral coeffi-

cients for each channel are then calculated by scaling the carrier signal using the respective scaling information for each channel.

[0013] As a result of this approach, only one set of spectral coefficients (*i.e.*, the coefficients of the carrier signal) needs to be transmitted, together with a small amount of side information (*i.e.*, the scaling information), instead of having to transmit a separate set of spectral components for each channel signal. For the two-channel stereo case, this results in a saving of almost 50% of the data rate for the intensity coded frequency regions.

[0014] Despite the advantages of this approach, however, excessive or uncontrolled application of the intensity stereo coding technique can lead to deterioration in the perceived stereo image, because the detailed structure of the signals over time is not preserved for time periods smaller than the granularity of the coding scheme (*e.g.*, 20ms per block). In particular, as a consequence of the use of a single carrier, all output signals which are reconstructed therefrom are necessarily scaled versions of each other. In other words, they have the same fine envelope structure for the duration of the coded block (*e.g.*, 10-20ms). This does not present a significant problem for stationary signals or for signals having similar fine envelope structures in the intensity stereo coded channels.

[0015] For transient signals with dissimilar envelopes in different channels, however, the original distribution of the envelope onsets between the coded channels cannot be recovered. For example, in a stereophonic recording of an applauding audience, the individual envelopes will be very different in the right and left channels due to the distinct clapping events happening at different times in both channels. Similar effects will occur for recordings by using stereophonic microphones, such that the spatial location of a sound source is, in essence, encoded as time differences or delays between the respective channel signals. Consequently, the stereo image quality of an intensity stereo coded/decoded signal will decrease significantly in these cases. The spatial impression tends to narrow, and the perceived stereo image tends to collapse into the center position. For critical signals, the achieved quality can no longer be considered acceptable.

[0016] Several strategies have been proposed in order to avoid deterioration in the stereo image of an intensity stereo encoded/decoded signal. Since using intensity stereo coding involves the risk of affecting the stereo image, it has been proposed to use the technique only in cases when the coder runs out of bits, so that severe quantization distortions, which would be perceived by the listener as being even more annoying, can be avoided. Alternatively, an algorithm can be employed which detects dissimilarities in the fine temporal structures of the channels. If a mismatch in envelopes is detected, intensity stereo coding is not applied in the given

block. Such an approach is described, for example, in "Intensity Stereo Coding" by J. Herre *et al.*, 96th Audio Engineering Society Convention, Amsterdam, February 1994. However, it is an obvious drawback of the prior proposed solutions that the potential for bit savings can no longer be fully exploited, given that the intensity stereo coding is disabled for such signals.

Summary of the Invention

[0017] In accordance with an illustrative embodiment of the present invention, the drawbacks of prior art techniques are overcome by a method and apparatus for performing joint stereo coding of multi-channel audio signals using intensity stereo coding techniques. In particular, predictive filtering techniques are applied to the spectral coefficient data, thereby preserving the fine time structure of the output signal of each channel, while maintaining the benefit of the high bit rate savings offered by intensity stereo coding. In one illustrative embodiment of the present invention, a method for enhancing the perceived stereo image of intensity stereo encoded/decoded signals is provided by applying the following processing steps in an encoder for two-channel stereophonic signals:

- The input signal of each channel is decomposed into spectral coefficients by a high-resolution filterbank/transform.
- Using a model, the one or more time-dependent masking thresholds are estimated for each channel. This advantageously gives the maximum coding error that can be introduced into the audio signal while still maintaining perceptually unimpaired signal quality.
- For each channel, a filter performing linear prediction in frequency is applied at the filterbank outputs, such that the residual, rather than the actual filterbank output signal, is used for the steps which follow.
- Intensity stereo coding techniques are applied for coding both residual signals into one carrier signal.
- The spectral values of the carrier signal are quantized and coded according to the precision corresponding to the masking threshold estimate(s).
- All relevant information (*i.e.*, the coded spectral values, intensity scaling data and prediction filter data for each channel, as well as additional side information) is packed into a bitstream and transmitted to the decoder.

[0018] Similarly, a decoder for joint stereo encoded signals, corresponding to the above-described illustrative

tive encoder and in accordance with another illustrative embodiment of the present invention, carries out the following processing steps:

- The bitstream is decoded and parsed into coded spectral data and side information.
- An inverse quantization of the quantized spectral values for the carrier signal is performed.
- Intensity stereo decoding is performed on the spectral values of the carrier signal, thereby producing (residual) signals for each channel.
- For each channel, inverse prediction filters, operating in frequency and corresponding to the prediction filters applied by the encoder used to encode the original signal, are applied to the residual signals.
- The values produced by the inverse prediction filters are mapped back into time domain representations using synthesis filterbanks.

Brief Description of the Drawings

[0019]

Fig. 1 shows a prior art encoder for two-channel stereophonic signals in which conventional intensity stereo coding techniques are employed.

Fig. 2 shows an encoder for two-channel stereophonic signals in accordance with an illustrative embodiment of the present invention.

Fig. 3 shows an illustrative implementation of the predictive filters of the illustrative encoder of Fig. 2.

Fig. 4 shows a prior art decoder for joint stereo coded signals, corresponding to the prior art encoder of Fig. 1, in which conventional intensity stereo coding techniques are employed.

Fig. 5 shows a decoder for joint stereo coded signals, corresponding to the illustrative encoder of Fig. 2, in accordance with an illustrative embodiment of the present invention.

Fig. 6 shows an illustrative implementation of the inverse predictive filters of the illustrative decoder of Fig. 5.

Fig. 7 shows a flow chart of a method of encoding two-channel stereophonic signals in accordance with an illustrative embodiment of the present invention.

Fig. 8 shows a flow chart of a method of decoding

joint stereo coded signals, corresponding to the illustrative encoding method shown in Fig. 7, in accordance with an illustrative embodiment of the present invention.

Detailed Description

Overview

[0020] The incorporation of a predictive filtering process into the encoder and decoder in accordance with certain illustrative embodiments of the present invention advantageously enhances the quality of the intensity stereo encoded/decoded signal by overcoming the limitation of prior art schemes whereby identical fine envelope structures are produced in all intensity stereo decoded channel signals. In particular, the illustrative encoding method overcomes the drawbacks of prior techniques by effectively extending the filterbank with the predictive filtering stage, such that the envelope information common over frequency is extracted as filter coefficients, and is, for the most part, stripped from the residual signal.

[0021] Specifically, for each input channel signal, a linear prediction is carried out on its corresponding spectral coefficient data, wherein the linear prediction is performed over frequency. Since predictive coding is applied to spectral domain data, the relations known for classical predictions are valid with the time and frequency domains interchanged. For example, the prediction error signal ideally has a "flat" (square of the) envelope, as opposed to having a "flat" power spectrum (a "pre-whitening" filter effect). The fine temporal structure information for each channel signal is contained in its prediction filter coefficients. Thus, it can be assumed that the carrier signal used for intensity stereo coding will also have a flat envelope, since it is generated by forming linear combinations of the (filtered) channel signals.

[0022] In a corresponding decoder in accordance with an illustrative embodiment of the present invention, each channel signal is re-scaled according to the transmitted scaling information, and the inverse filtering process is applied to the spectral coefficients. In this way, the inverse "pre-whitening" process is performed on the envelope of each decoded channel signal, effectively reintroducing the envelope information into the spectral coefficients. Since this is done individually for each channel, the extended encoding/decoding system is capable of reproducing different individual fine envelope structures for each channel signal. Note that, in effect, using a combination of filterbank and linear prediction in frequency is equivalent to using an adaptive filterbank matched to the envelope of the input signal. Since the process of envelope shaping a signal can be performed either for the entire spectrum of the signal or for only part thereof, this time-domain envelope control can be advantageously applied in any necessary frequency-dependent fashion.

[0023] And in accordance with another embodiment of the present invention, the bitstream which is, for example, by the illustrative encoder described above (and described in further detail below with reference to Figs. 2, 3 and 7) may be advantageously stored on a storage medium such as a Compact Disc or a Digital Audio Tape, or stored in a semiconductor memory device. Such a storage medium may then be "read back" to supply the bitstream for subsequent decoding by, for example, the illustrative decoder described above (and described in further detail below with reference to Figs. 5, 6 and 8). In this manner, a substantial quantity of audio data (e.g., music) may be compressed onto the given storage medium without loss of (perceptual) quality in the reconstructed signal.

A prior art encoder

[0024] Fig. 1 shows a prior art perceptual encoder for two-channel stereophonic signals in which conventional intensity stereo coding techniques are employed. The encoder of Fig. 1 operates as follows:

- The left and right input signals, $x_l(k)$ and $x_r(k)$, are each individually decomposed into spectral coefficients by analysis filterbank/transform modules 12l and 12r, respectively, resulting in corresponding sets of "n" spectral components, $y_l(b,0 \dots n-1)$ and $y_r(b,0 \dots n-1)$, respectively, for each analysis block b. where "n" is the number of spectral coefficients per analysis block (i.e., the block size). Each spectral component $y_l(b,i)$ or $y_r(b,i)$ is associated with an analysis frequency in accordance with the particular filterbank employed.
- For each channel, perceptual model 11l or 11r estimates the required coding precision for perceptually transparent quality of the encoded/decoded signal. This estimation data may, for example, be based on the minimum signal-to-noise ratio (SNR) required in each coder band, and is passed to the quantization/encoding module.
- The spectral values for both the left and the right channel, $y_l(b,0 \dots n-1)$ and $y_r(b,0 \dots n-1)$, are provided to intensity stereo encoding module 13, which performs conventional intensity stereo encoding. For portions of the spectrum which are to be excluded from intensity stereo coding, the corresponding values of $y_l(b,0 \dots n-1)$ and $y_r(b,0 \dots n-1)$ may be passed directly to the quantization and coding stage. For portions of the spectrum which are to make use of intensity stereo coding (i.e., preferably the high-frequency portions thereof), the intensity stereo coding process is performed as follows. From each of the signals $y_l()$ and $y_r()$, scaling information is extracted for each coder frequency band (e.g., peak amplitude or total energy), and a single

carrier signal $y_i()$ is generated by combining the corresponding $y_l()$ and $y_r()$ values. Thus, for spectral portions coded in intensity stereo, only one set of values $y_i()$ for both channels, plus scaling side information for each channel, is provided to the quantization and coding stage. Alternatively, combined scaling information for both channels together with directional information can be used (along with the single carrier signal).

- The spectral components at the output of the intensity stereo encoding stage, consisting of separate values $y_l()$ and $y_r()$ and common values $y_i()$, are quantized and mapped to transmission symbols by quantization and encoding module 14. This module takes into account the required coding precision as determined by perceptual models 11l and 11r.
- The transmission symbol values generated by quantization and encoding module 14, together with further side information, are passed to bitstream encoder/multiplexer 15 and are thereby transmitted in the encoded bitstream. For coder frequency bands which use intensity stereo coding, the scaling information delivered by intensity stereo encoding module 13 is also provided to bitstream encoder/multiplexer 15 and thereby transmitted in the encoded bitstream as well.

An illustrative encoder

[0025] Fig. 2 shows an encoder for two-channel stereophonic signals in accordance with an illustrative embodiment of the present invention. The operation of the illustrative encoder of Fig. 2 is similar to that of the prior art encoder shown in Fig. 1, except that, for each channel, a predictive filtering stage is introduced between the corresponding analysis filterbank and the intensity stereo encoding module. That is, predictive filters 16l and 16r are applied to the outputs of analysis filterbanks 12l and 12r, respectively. As such, the spectral values, $y_l(b,0 \dots n-1)$ and $y_r(b,0 \dots n-1)$, are replaced by the output values of the predictive filtering process, $y'_l(b,0 \dots n-1)$ and $y'_r(b,0 \dots n-1)$, respectively, before being provided to intensity stereo encoding module 13.

[0026] Fig. 3 shows an illustrative implementation of the predictive filters of the illustrative encoder of Fig. 2. Specifically, inside the predictive filtering stage for each channel, a linear prediction is performed across frequency (as opposed, for example, to predictive coding which is performed across time, such as is employed by subband-ADPCM coders). To this end "rotating switch" 43 operates to bring spectral values $y(b,0 \dots n-1)$ into a serial order prior to processing, and the resulting output values $y'(b,0 \dots n-1)$ are provided in parallel thereafter by "rotating switch" 46. (Note that the use of "rotating switches" as a mechanism for conversion between serial and parallel orderings is used herein only for the pur-

pose of convenience and ease of understanding. As will be obvious to those of ordinary skill in the art, no such physical switching device need be provided. Rather, conversions between serial and parallel orderings may be performed in any of a number of conventional ways familiar to those skilled in the art, including by the use of software alone.) Although the illustrative embodiment shown herein performs the processing of the spectral values in order of increasing frequency, alternative embodiments may, for example, perform the processing thereof in order of decreasing frequency. Other orderings are also possible, as would be clear to one of ordinary skill in the art.

[0027] Specifically, as can be seen from the figure, the resultant output values, $y'(b.0...n-1)$, are computed from the input values, $y(b.0...n-1)$, by subtracting (with use of subtractor 48) the predicted value (predicted by predictor 47) from the input values, so that only the prediction error signal is passed on. Note that the combination of predictor 47 and subtractor 48, labelled in the figure as envelope pre-whitening filter 44, functions to equalize the temporal shape of the corresponding time signal.

[0028] The process performed by predictive filters 161 and 16r of the illustrative encoder of Fig. 2 can be performed either for the entire spectrum (*i.e.*, for all spectral coefficients), or, alternatively, for only a portion of the spectrum (*i.e.*, a subset of the spectral coefficients). Moreover, different predictor filters (*e.g.*, different predictors 47 as shown in Fig. 3) can be used for different portions of the signal spectrum. In this manner, the above-described method for time-domain envelope control can be applied in any necessary frequency-dependent fashion.

[0029] In order to enable the proper decoding of the signal the bitstream advantageously includes certain additional side information. For example, one field of such information might indicate the use of predictive filtering and if applicable, the number of different prediction filters. If predictive filtering is used, additional fields in the bitstream may be transmitted for each prediction filter indicating the target frequency range of the respective filter and its filter coefficients. Thus, as shown in Fig. 2 by the dashed lines labelled "L Filter Data" and "R Filter Data," predictive filters 161 and 16r provide the necessary information to bitstream encoder/multiplexer 17 for inclusion in the transmitted bitstream.

[0030] Fig. 7 shows a flow chart of a method of encoding two-channel stereophonic signals in accordance with an illustrative embodiment of the present invention. The illustrative example shown in this flow chart implements certain relevant portions of the illustrative encoder of Fig. 2. Specifically, the flow chart shows the front-end portion of the encoder for a single one of the channels, including the envelope pre-whitening process using a single prediction filter. This pre-whitening process is carried out after the calculation of the spectral values by the analysis filterbank, as shown in step 61 of the figure.

[0031] Specifically, after the analysis filterbank is run, the order of the prediction filter is set and the target frequency range is defined (step 62). These parameters may illustratively be set to a filter order of 15 and a target frequency range comprising the entire frequency range that will be coded using intensity stereo coding (*e.g.*, from 4 kHz to 20 kHz). In this manner, the scheme is advantageously configured to provide one set of individual fine temporal structure data for each audio channel. In step 63, the prediction filter is determined by using the range of spectral coefficients matching the target frequency range, and by applying a conventional method for predictive coding as is well known, for example, in the context of Differential Pulse Code Modulation (DPCM) coders. For example, the autocorrelation function of the coefficients may be calculated and used in a conventional Levinson-Durbin recursion algorithm, well known to those skilled in the art. As a result, the predictor filter coefficients, the corresponding reflection coefficients ("PARCOR" coefficients), and the expected prediction gain are known.

[0032] If the expected prediction gain exceeds a certain threshold (*e.g.*, 2 dB), as determined by decision 64, the predictive filtering procedure of steps 65 through 67 is used. In this case, the prediction filter coefficients are quantized (in step 65) as required for transmission to the decoder as part of the side information. Then, in step 66, the prediction filter is applied to the range of spectral coefficients matching the target frequency range where the quantized filter coefficients are used. For all further processing, therefore, the spectral coefficients are replaced by the output of the filtering process. Finally, in step 67, a field of the bitstream to be transmitted is set to indicate the use of predictive filtering ("prediction flag" on). In addition, the target frequency range, the order of the prediction filter, and information describing its filter coefficients are also included in the bitstream.

[0033] If, on the other hand, the expected prediction gain does not exceed the decision threshold as determined by decision 64, step 68 sets a field in the bitstream to indicate that no predictive filtering has been used ("prediction flag" off). Finally, after the above-described processing is complete, conventional steps as performed in prior art encoders (such as those carried out by the encoder of Fig. 1) are performed -- that is, the intensity stereo encoding process is applied to the spectral coefficients (which may now be residual data), the results of the intensity stereo encoding process are quantized and encoded, and the actual bitstream to be transmitted is encoded for transmission (with the appropriate side information multiplexed therein). Note, however, that bitstream encoder/multiplexer 17 of the illustrative encoder of Fig. 2 replaces conventional bitstream encoder/multiplexer 15 of the prior art encoder of Fig. 1. so that the additional side information provided by predictive filters 161 and 16r (*i.e.*, "L Filter Data" and "R Filter Data") may be advantageously encoded and

transmitted in the resultant bitstream.

A prior art decoder

[0034] Fig. 4 shows a prior art decoder for joint stereo coded signals, corresponding to the prior art encoder of Fig. 1, in which conventional intensity stereo coding techniques are employed. Specifically, the decoder of Fig. 4 performs the following steps:

- The incoming bitstream is parsed by bitstream decoder/demultiplexer 21, and the transmission symbols for the spectral coefficients are passed on to decoding and inverse quantization module 22, together with the quantization related side information.
- In decoding and inverse quantization module 22, the quantized spectral values, $y_{ql}()$, $y_{qr}()$ and $y_{qi}()$, are reconstructed. These signals correspond to the independently coded left channel signal portion, the independently coded right channel signal portion, and the intensity stereo carrier signal, respectively.
- From the reconstructed spectral values of the carrier signal and the transmitted scaling information, the missing portions of the $y_{ql}()$ and $y_{qr}()$ spectra for the left and right channel signals are calculated with use of a conventional intensity stereo decoding process, which is performed by intensity stereo decoding module 23. At the output of this module, two complete (and independent) channel spectral signals, $y_{ql}()$ and $y_{qr}()$, corresponding to the left and right channels, respectively, are available.
- Finally, each of the left and right channel spectral signals, $y_{ql}()$ and $y_{qr}()$, are mapped back into a time domain representation by synthesis filterbanks 24l and 24r, respectively, thereby resulting in the final output signals $x_l'(k)$ and $x_r'(k)$.

An illustrative decoder

[0035] Fig. 5 shows a decoder for joint stereo coded signals, corresponding to the illustrative encoder of Fig. 2, in accordance with an illustrative embodiment of the present invention. The operation of the illustrative decoder of Fig. 5 is similar to that of the prior art decoder shown in Fig. 4, except that, for each channel, an inverse predictive filtering stage is introduced between the intensity stereo decoding and the corresponding synthesis filterbanks. That is, inverse predictive filters 26l and 26r are inserted prior to synthesis filterbanks 24l and 24r, respectively. Thus, the spectral values, $y_{ql}()$ and $y_{qr}()$, as generated by intensity stereo decoding module 23, are replaced by the output values of the corresponding inverse predictive filtering processes, $y_{ql}'()$ and $y_{qr}'()$, respectively, before being provided to their

corresponding synthesis filterbanks (synthesis filterbanks 24l and 24r).

[0036] Fig. 6 shows an illustrative implementation of the inverse predictive filters of the illustrative decoder of Fig. 5. Specifically, within the inverse predictive filters, a linear filtering operation is performed across (as opposed to performing predictive coding across time as in subband-ADPCM coders). In a similar manner to that shown in the prediction filter implementation of Fig. 3, "rotating switch" 33 of Fig. 6 is used to bring the spectral values $y_q(b,0 \dots n-1)$ into a serial order prior to processing, and "rotating switch" 36 of the figure is used to bring the resulting output values $y_q'(b,0 \dots n-1)$ into a parallel order thereafter. (Once again, note that the use of "rotating switches" as a mechanism for conversion between serial and parallel orderings is provided herein only for the purpose of convenience and ease of understanding. As will be obvious to those of ordinary skill in the art, no such physical switching device need be provided. Rather, conversions between serial and parallel orderings may be performed in any of a number of conventional ways familiar to those skilled in the art, including by the use of software alone.) Again, as in the case of the illustrative encoder described above, processing in order of increasing or decreasing frequency is possible, as well as other possible orderings obvious to those skilled in the art.

[0037] Specifically, as can be seen from the figure, the output values, $y_q'(b,0 \dots n-1)$, are computed from the input values, $y_q(b,0 \dots n-1)$, by applying the inverse of the envelope pre-whitening filter used in the corresponding encoder. In particular, the output values are computed from the input values by adding (with use of adder 38) the predicted values (predicted by predictor 37) to the input values as shown. Note that the combination of predictor 37 and adder 38, labelled in the figure as envelope shaping filter 34, functions to re-introduce the temporal shape of the original time signal.

[0038] As described above in the discussion of the illustrative encoder of Figs. 2 and 3, the above-described filtering process can be performed either for the entire spectrum (*i.e.*, for all spectral coefficients), or for only a portion of the spectrum (*i.e.*, a subset of the spectral coefficients). Moreover, different predictor filters (*e.g.*, different predictors 37 as shown in Fig. 6) can be used for different parts of the signal spectrum. In such a case (in order to execute the proper decoding of the signal), the illustrative decoder of Fig. 5 advantageously decodes from the bitstream the additional side information (labelled in the figure as "L Filter Data" and "R Filter Data") which had been transmitted by the encoder, and supplies this data to inverse predictive filters 26l and 26r. In this manner, predicative decoding can be applied in each specified target frequency range with a corresponding prediction filter.

[0039] Fig. 8 shows a flow chart of a method of decoding joint stereo coded signals, corresponding to the illustrative encoding method shown in Fig. 7, in accord-

ance with an illustrative embodiment of the present invention. The illustrative example shown in this flow chart implements certain relevant portions of the illustrative decoder of Fig. 5. Specifically, the flow chart shows the back-end portion of the decoder for a single one of the channels, including the envelope shaping process using a single (inverse) prediction filter. The processing which is performed by the decoder prior to those steps shown in the flow chart of Fig. 8 comprises conventional steps performed in prior art decoders (such as those carried out by the decoder of Fig. 4) -- that is, the bitstream is decoded/demultiplexed, the resultant data is decoded and inverse quantized, and the intensity stereo decoding process is performed. Note, however, that bitstream decoder/demultiplexer 25 of the illustrative decoder of Fig. 5 replaces conventional bitstream decoder/demultiplexer 21 of the prior art decoder of Fig. 4, so that the additional side information provided by the encoder (e.g., "L Filter Data" and "R Filter Data") may be advantageously decoded and provided to inverse predictive filters 26l and 26r.

[0040] After the intensity stereo decoding has been completed, the data from the bitstream which signals the use of predictive filtering is checked (by decision 72). If the data indicates that predictive filtering was performed in the encoder (*i.e.*, the "prediction flag" is on), then the extended decoding process of steps 73 and 74 is carried out. Specifically, the target frequency range of the prediction filtering, the order of the pre-whitening (prediction) filter, and information describing the coefficients of the filter are retrieved from the (previously decoded) side information (step 73). Then, the inverse (decoder) prediction filter (*i.e.*, the envelope shaping filter) is applied to the range of spectral coefficients matching the target frequency range (step 74). In either case (*i.e.*, whether predictive filtering was performed or not), the decoder processing completes by running the synthesis filterbank (for each channel) from the spectral coefficients (as processed by the envelope shaping filter, if applicable), as shown in step 75.

Conclusion

[0041] Using the above-described process in accordance with the illustrative embodiments of the present invention (*i.e.*, predictive filtering in the encoder and inverse filtering in the decoder), a straightforward envelope shaping effect can be achieved for certain conventional block transforms including the Discrete Fourier Transform (DFT) or the Discrete Cosine Transform (DCT), both well-known to those of ordinary skill in the art. If, for example, a perceptual coder in accordance with the present invention uses a critically subsampled filterbank with overlapping windows -- *e.g.*, a conventional Modified Discrete Cosine Transform or another conventional filterbank based on Time Domain Aliasing Cancellation (TDAC) -- the resultant envelope shaping effect is subject to the time domain aliasing effects in-

herent in the filterbank. For example, in the case of a MDCT, one mirroring (*i.e.*, aliasing) operation per window half takes place, and the fine envelope structure appears mirrored (*i.e.*, aliased) within the left and the right half of the window after decoding, respectively. Since the final filterbank output is obtained by applying a synthesis window to the output of each inverse transform and performing an overlap-add of these data segments, the undesired aliased components are attenuated depending on the synthesis window used. Thus, it is advantageous to choose a filterbank window that exhibits only a small overlap between subsequent blocks, so that the temporal aliasing effect is minimized. An appropriate strategy in the encoder can, for example, adaptively select a window with a low degree of overlap for critical signals, thereby providing improved frequency selectivity. The implementation details of such a strategy will be obvious to those skilled in the art.

[0042] Although a number of specific embodiments of this invention have been shown and described herein, it is to be understood that these embodiments are merely illustrative of the many possible specific arrangements which can be devised in application of the principles of the invention. For example, although the illustrative embodiments which have been shown and described herein have been limited to the encoding and decoding of stereophonic audio signals comprising only two channels, alternative embodiments which may be used for the encoding and decoding of stereophonic audio signals having more than two channels will be obvious to those of ordinary skill in the art based on the disclosure provided herein. In addition, numerous and varied other arrangements can be devised in accordance with these principles by those of ordinary skill in the art without departing from the scope of the invention.

Claims

1. A method of performing joint stereo coding of a multichannel audio signal to generate an encoded signal, the method comprising the steps of:
 - (a) performing a spectral decomposition of a first audio channel signal into a plurality of first spectral component signals;
 - (b) generating a first prediction signal representative of a prediction of one of said first spectral component signals, said prediction based on one or more other ones of said first spectral component signals;
 - (c) comparing the first prediction signal with said one of said first spectral component signals to generate a first prediction error signal;
 - (d) performing a spectral decomposition of a second audio channel signal into a plurality of second spectral component signals;
 - (e) performing joint stereo coding of said one

- of said first spectral component signals and one of said second spectral component signals to generate a jointly coded spectral component signal, said coding based on the first prediction error signal; and
 (f) generating the encoded signal based on the jointly coded spectral component signal. 5
2. The method of claim 1 further comprising the steps of: 10
- (g) generating a second prediction signal representative of a prediction of said one of said second spectral component signals, said prediction based on one or more other ones of said second spectral component signals; and 15
- (h) comparing the second prediction signal with said one of said second spectral component signals to generate a second prediction error signal; 20
- and wherein the step of performing joint stereo coding of said one of said first spectral component signals and said one of said second spectral component signals is further based on said second prediction error signal. 25
3. The method of claim 1 wherein the step of performing joint stereo coding of said one of said first spectral component signals and said one of said second spectral component signals comprises performing intensity stereo coding of said one of said first spectral component signals and said one of said second spectral component signals. 30
4. The method of claim 1 wherein the step of generating the encoded signal based on the jointly coded spectral component signal comprises quantizing the jointly coded spectral component signal. 40
5. The method of claim 4 wherein said quantization of the jointly coded spectral component signal is based on a perceptual model.
6. A method as claimed in any of the preceding claims further including the step of storing said encoded signal on a storage medium. 45
7. The method of claim 6 wherein the storage medium comprises a compact disc. 50
8. The method of claim 6 wherein the storage medium comprises a digital audio tape.
9. The method of claim 6 wherein the storage medium comprises a semiconductor memory. 55
10. A method of decoding an encoded signal to generate a reconstructed multichannel audio signal, the encoded signal comprising a joint stereo coding of an original multichannel audio signal, the method comprising the steps of:
- (a) performing joint stereo decoding of the encoded signal to generate a plurality of decoded channel signals, each decoded channel signal comprising a plurality of decoded spectral component prediction error signals;
- (b) generating a first spectral component signal based on one or more of said spectral component prediction error signals comprised in a first one of said decoded channel signals;
- (c) generating a first prediction signal representative of a prediction of a second spectral component signal, said prediction based on said first spectral component signal;
- (d) generating the second spectral component signal based on the first prediction signal and on one or more of said spectral component prediction error signals comprised in the first one of said decoded channel signals; and
- (e) generating a first channel of the reconstructed multi-channel audio signal based on the first and second spectral component signals.
11. The method of claim 10 further comprising the steps of:
- (f) generating a third spectral component signal based on one or more of said spectral component prediction error signals comprised in a second one of said decoded channel signals;
- (g) generating a second prediction signal representative of a prediction of a fourth spectral component signal, said prediction based on said third spectral component signal;
- (h) generating the fourth spectral component signal based on the second prediction signal and on one or more of said spectral component prediction error signals comprised in the second one of said decoded channel signals; and
- (i) generating a second channel of the reconstructed multichannel audio signal based on the third and fourth spectral component signals.
12. The method of claim 10 wherein the step of performing joint stereo decoding of the encoded signal comprises performing intensity stereo decoding of the encoded signal.
13. An encoder for performing joint stereo coding of a multi-channel audio signal to generate an encoded signal, the encoder comprising:
- (a) a first filterbank (12l, 12r) which performs a spectral decomposition of a first audio channel

signal into a plurality of first spectral component signals;

(b) a first prediction filter (16l, 16r, 47) which generates a first prediction signal representative of a prediction of one of said first spectral component signals, said prediction filter responsive to one or more other ones of said first spectral component signals;

(c) a first comparator (48) which compares the first prediction signal with said one of said first spectral component signals to generate a first prediction error signal;

(d) a second filterbank (12l, 12r) which performs a spectral decomposition of a second audio channel signal into a plurality of second spectral component signals;

(e) a joint stereo coder (13) which performs joint stereo coding of said one of said first spectral component signals and one of said second spectral component signals to generate a jointly coded spectral component signal, said coding based on the first prediction error signal; and

(f) a coder (14) which generates the encoded signal based on the jointly coded spectral component signal.

14. The encoder of claim 13 further comprising:

(g) a second prediction filter (16l, 16r, 47) which generates a second prediction signal representative of a prediction of said one of said second spectral component signals, said prediction based on one or more other ones of said second spectral component signals; and

(h) a second comparator (48) which compares the second prediction signal with said one of said second spectral component signals to generate a second prediction error signal;

and wherein the joint stereo coder performs joint stereo coding further based on said second prediction error signal.

15. The encoder of claim 13 wherein the joint stereo coder comprises an intensity stereo coder which performs intensity stereo coding of said one of said first spectral component signals and said one of said second spectral component signals.

16. The encoder of claim 13 wherein the coder which generates the encoded signal based on the jointly coded spectral component signal comprises a quantizer which quantizes the jointly coded spectral component signal.

17. The encoder of claim 16 wherein the quantizer is based on a perceptual model.

18. A decoder for decoding an encoded signal to generate a reconstructed multichannel audio signal, the encoded signal comprising a joint stereo coding of an original multichannel audio signal, the method comprising:

(a) a joint stereo decoder (23) which performs joint stereo decoding of the encoded signal to generate a plurality of decoded channel signals, each decoded channel signal comprising a plurality of decoded spectral component prediction error signals;

(b) means for generating a first spectral component signal based on one or more of said spectral component prediction error signals comprised in a first one of said decoded channel signals;

(c) a first prediction filter (26l, 26r) which generates a first prediction signal representative of a prediction of a second spectral component signal, said prediction based on said first spectral component signal;

(d) means for generating the second spectral component signal based on the first prediction signal and on one or more of said spectral component prediction error signals comprised in the first one of said decoded channel signals; and

(e) a first filterbank (24l, 24r) which generates a first channel of the reconstructed multichannel audio signal based on the first and second spectral component signals.

19. The decoder of claim 18 further comprising:

(f) means for generating a third spectral component signal based on one or more of said spectral component prediction error signals comprised in a second one of said decoded channel signals;

(g) a second prediction filter which generates a second prediction signal representative of a prediction of a fourth spectral component signal, said prediction based on said third spectral component signal;

(h) means for generating the fourth spectral component signal based on the second prediction signal and on one or more of said spectral component prediction error signals comprised in the second one of said decoded channel signals; and

(i) a second filterbank which generates a second channel of the reconstructed multichannel audio signal based on the third and fourth spectral component signals.

20. The decoder of claim 18 wherein the joint stereo decoder comprises an intensity stereo decoder which performs intensity stereo decoding of the encoded

signal.

Patentansprüche

1. Verfahren zum Kombinationsstereokodieren eines Mehrkanal-Audiosignals, um ein kodiertes Signal zu erzeugen, wobei das Verfahren die folgenden Schritte aufweist:

(a) Ausführen einer Spektralzerlegung eines ersten Audiokanalsignals in eine Mehrzahl an ersten Spektralkomponentensignalen;
 (b) Erzeugen eines ersten Prädiktionssignals, das eine Vorhersage eines der ersten Spektralkomponentensignale darstellt, wobei die Vorhersage auf einem oder mehreren weiteren der ersten Spektralkomponentensignale beruht;
 (c) Vergleichen des ersten Prädiktionssignals mit jenem einen der ersten Spektralkomponentensignale, um ein erstes Prädiktionsfehlersignal zu erzeugen;
 (d) Ausführen einer Spektralzerlegung eines zweiten Audiokanalsignals in eine Mehrzahl an zweiten Spektralkomponentensignalen;
 (e) Kombinationsstereokodieren jenes einen der ersten Spektralkomponentensignale sowie eines der zweiten Spektralkomponentensignale, um ein kombinationskodiertes Spektralkomponentensignal zu erzeugen, wobei die Kodierung auf dem ersten Prädiktionsfehlersignal beruht; und
 (f) Erzeugen des kodierten Signals auf der Grundlage des kombinationskodierten Spektralkomponentensignals.

2. Verfahren nach Anspruch 1, welches außerdem die folgenden Schritte aufweist:

(g) Erzeugen eines zweiten Prädiktionssignals, das eine Vorhersage jenes einen der zweiten Spektralkomponentensignale darstellt, wobei die Vorhersage auf einem oder mehreren weiteren der zweiten Spektralkomponentensignale beruht; und
 (h) Vergleichen des zweiten Prädiktionssignals mit jenem einen der zweiten Spektralkomponentensignale, um ein zweites Prädiktionsfehlersignal zu erzeugen;

und wobei der Schritt des Kombinationsstereokodierens jenes einen der ersten Spektralkomponentensignale sowie jenes einen der zweiten Spektralkomponentensignale außerdem auf dem zweiten Prädiktionsfehlersignal beruht.

3. Verfahren nach Anspruch 1, wobei der Schritt des Kombinationsstereokodierens jenes einen der er-

sten Spektralkomponentensignale sowie jenes einen der zweiten Spektralkomponentensignale das Intensitätsstereokodieren jenes einen der ersten Spektralkomponentensignale sowie jenes einen der zweiten Spektralkomponentensignale umfasst.

4. Verfahren nach Anspruch 1, wobei der Schritt des Erzeugens des kodierten Signals auf der Grundlage des kombinationskodierten Spektralkomponentensignals das Quantisieren des kombinationskodierten Spektralkomponentensignals umfasst.
 5. Verfahren nach Anspruch 4, wobei die Quantisierung des kombinationskodierten Spektralkomponentensignals auf einem Wahrnehmungsmodell beruht.
 6. Verfahren nach einem der vorhergehenden Ansprüche, das weiterhin den Schritt des Speicherns des kodierten Signals auf einem Speichermedium umfasst.
 7. Verfahren nach Anspruch 6, wobei das Speichermedium eine Compact Disc umfasst.
 8. Verfahren nach Anspruch 6, wobei das Speichermedium ein Digitaltonband umfasst.
 9. Verfahren nach Anspruch 6, wobei das Speichermedium einen Halbleiterspeicher umfasst.
 10. Verfahren zum Dekodieren eines kodierten Signals, um ein rekonstruiertes Mehrkanal-Audiosignal zu erzeugen, wobei das kodierte Signal eine Kombinationsstereokodierung eines ursprünglichen Mehrkanal-Audiosignals umfasst, und wobei das Verfahren die folgenden Schritte aufweist:

(a) Kombinationsstereodekodieren des kodierten Signals, um eine Mehrzahl an dekodierten Kanalsignalen zu erzeugen, wobei jedes dekodierte Kanalsignal eine Mehrzahl an dekodierten Spektralkomponenten-Prädiktionsfehlersignalen umfasst;
 (b) Erzeugen eines ersten Spektralkomponentensignals auf der Grundlage eines oder mehrerer dieser Spektralkomponenten-Prädiktionsfehlersignale, die in einem ersten der dekodierten Kanalsignale inbegriffen sind;
 (c) Erzeugen eines ersten Prädiktionssignals, das eine Vorhersage eines zweiten Spektralkomponentensignals darstellt, wobei die Vorhersage auf dem ersten Spektralkomponentensignal beruht;
 (d) Erzeugen des zweiten Spektralkomponentensignals auf der Grundlage des ersten Prädiktionssignals sowie eines oder mehrerer der Spektralkomponenten-Prädiktionsfehlersigna-

le, die in dem ersten der dekodierten Kanalsignale inbegriffen sind; und

(e) Erzeugen eines ersten Kanals des rekonstruierten Mehrkanal-Audiosignals auf der Grundlage der ersten und zweiten Spektralkomponentensignale.

11. Verfahren nach Anspruch 10, welches außerdem die folgenden Schritte aufweist:

(f) Erzeugen eines dritten Spektralkomponentensignals auf der Grundlage eines oder mehrerer der Spektralkomponenten-Prädiktionsfehlersignale, die in einem zweiten der dekodierten Kanalsignale inbegriffen sind;

(g) Erzeugen eines zweiten Prädiktionssignals, das eine Vorhersage eines vierten Spektralkomponentensignals darstellt, wobei die Vorhersage auf dem dritten Spektralkomponentensignal beruht;

(h) Erzeugen des vierten Spektralkomponentensignals auf der Grundlage des zweiten Prädiktionssignals sowie eines oder mehrerer der Spektralkomponenten-Prädiktionsfehlersignale, die in dem zweiten der dekodierten Kanalsignale inbegriffen sind; und

(i) Erzeugen eines zweiten Kanals des rekonstruierten Mehrkanal-Audiosignals auf der Grundlage der dritten und vierten Spektralkomponentensignale.

12. Verfahren nach Anspruch 10, wobei der Schritt des Kombinationsstereodekodierens des kodierten Signals das Intensitätsstereodekodieren des kodierten Signals umfasst.

13. Kodierer zum Kombinationsstereokodieren eines Mehrkanal-Audiosignals, um ein kodiertes Signal zu erzeugen, wobei der Kodierer Folgendes aufweist:

(a) eine erste Filterbank (121, 12r), die eine Spektralzerlegung eines ersten Audiokanalsignals in eine Mehrzahl an ersten Spektralkomponentensignalen ausführt;

(b) einen ersten Prädiktionsfilter (161, 16r, 47), der ein erstes Prädiktionsignal erzeugt, welches eine Vorhersage eines der ersten Spektralkomponentensignale darstellt, wobei der Prädiktionsfilter auf eines oder mehrere weitere der ersten Spektralkomponentensignale anspricht;

(c) einen ersten Vergleich (48), der das erste Prädiktionsignal mit jenem einen der ersten Spektralkomponentensignale vergleicht, um ein erstes Prädiktionsfehlersignal zu erzeugen;

(d) eine zweite Filterbank (121, 12r), die eine Spektralzerlegung eines zweiten Audiokanalsignals

in eine Mehrzahl an zweiten Spektralkomponentensignalen ausführt;

(e) einen Kombinationsstereokodierer (13), der das Kombinationsstereokodieren jenes einen der ersten Spektralkomponentensignale sowie eines der zweiten Spektralkomponentensignale ausführt, um ein kombinationskodiertes Spektralkomponentensignal zu erzeugen, wobei die Kodierung auf dem ersten Prädiktionsfehlersignal beruht; und

(f) einen Kodierer (14), der das kodierte Signal auf der Grundlage des kombinationskodierten Spektralkomponentensignals erzeugt.

14. Kodierer nach Anspruch 13, welcher außerdem Folgendes umfasst:

(g) einen zweiten Prädiktionsfilter (161, 16r, 47), der ein zweites Prädiktionsignal erzeugt, welches eine Vorhersage jenes einen der zweiten Spektralkomponentensignale darstellt, wobei die Vorhersage auf einem oder mehreren weiteren der zweiten Spektralkomponentensignale beruht; und

(h) einen zweiten Vergleich (48), der das zweite Prädiktionsignal mit jenem einen der zweiten Spektralkomponentensignale vergleicht, um ein zweites Prädiktionsfehlersignal zu erzeugen;

und wobei der Kombinationsstereokodierer das Kombinationsstereokodieren außerdem auf der Grundlage des zweiten Prädiktionsfehlersignals ausführt.

15. Kodierer nach Anspruch 13, wobei der Kombinationsstereokodierer einen Intensitätsstereokodierer umfasst, welcher das Intensitätsstereokodieren jenes einen der ersten Spektralkomponentensignale sowie jenes einen der zweiten Spektralkomponentensignale ausführt.

16. Kodierer nach Anspruch 13, wobei der Kodierer, der das kodierte Signal auf der Grundlage des kombinationskodierten Spektralkomponentensignals erzeugt, einen Quantisierer umfasst, der das kombinationskodierte Spektralkomponentensignal quantisiert.

17. Kodierer nach Anspruch 16, wobei der Quantisierer auf einem Wahrnehmungsmodell beruht.

18. Dekodierer zum Dekodieren eines kodierten Signals, um ein rekonstruiertes Mehrkanal-Audiosignal zu erzeugen, wobei das kodierte Signal ein Kombinationsstereokodieren eines ursprünglichen Mehrkanal-Audiosignals umfasst, und wobei das Verfahren Folgendes aufweist:

(a) einen Kombinationsstereodekodierer (23), der das Kombinationsstereodekodieren der kodierten Signale ausführt, um eine Mehrzahl an dekodierten Kanalsignalen zu erzeugen, wobei jedes dekodierte Kanalsignal eine Mehrzahl an dekodierten Spektralkomponenten-Prädiktionsfehlersignalen umfasst;

(b) Mittel zum Erzeugen eines ersten Spektralkomponentensignals auf der Grundlage eines oder mehrerer der Spektralkomponenten-Prädiktionsfehlersignale, die in einem ersten der dekodierten Kanalsignale inbegriffen sind;

(c) einen ersten Prädiktionsfilter (261, 26r), der ein erstes Prädiktionssignal erzeugt, das eine Vorhersage eines zweiten Spektralkomponentensignals darstellt, wobei die Vorhersage auf dem ersten Spektralkomponentensignal beruht;

(d) Mittel zum Erzeugen des zweiten Spektralkomponentensignals auf der Grundlage des ersten Prädiktionssignals sowie eines oder mehrerer der Spektralkomponenten-Prädiktionsfehlersignale, die in dem ersten der dekodierten Kanalsignale inbegriffen sind; und

(e) eine erste Filterbank (241, 24r), die einen ersten Kanal des rekonstruierten Mehrkanal-Audiosignals auf der Grundlage der ersten und zweiten Spektralkomponentensignale erzeugt.

19. Dekodierer nach Anspruch 18, welcher außerdem Folgendes umfasst:

(f) Mittel zum Erzeugen eines dritten Spektralkomponentensignals auf der Grundlage eines oder mehrerer der Spektralkomponenten-Prädiktionsfehlersignale, die in einem zweiten der dekodierten Kanalsignale inbegriffen sind; .

(g) einen zweiten Prädiktionsfilter, welcher ein zweites Prädiktionssignal erzeugt, das eine Vorhersage eines vierten Spektralkomponentensignals darstellt, wobei die Vorhersage auf dem dritten Spektralkomponentensignal beruht;

(h) Mittel zum Erzeugen des vierten Spektralkomponentensignals auf der Grundlage des zweiten Prädiktionssignals sowie eines oder mehrerer der Spektralkomponenten-Prädiktionsfehlersignale, die in dem zweiten der dekodierten Kanalsignale inbegriffen sind; und

(i) eine zweite Filterbank, die einen zweiten Kanal des rekonstruierten Mehrkanal-Audiosignals auf der Grundlage der dritten und vierten Spektralkomponentensignale erzeugt.

20. Dekodierer nach Anspruch 18, wobei der Kombinationsstereodekodierer einen Intensitätsstereodekodierer umfasst, welcher das Intensitätsstereodekodieren des kodierten Signals ausführt.

Revendications

1. Procédé d'exécution d'un codage stéréo combiné d'un signal audio multicanaux afin de générer un signal codé, le procédé comprenant les étapes suivantes:

(a) exécution d'une décomposition spectrale d'un premier signal de canal audio en une pluralité de premiers signaux de composantes spectrales;

(b) génération d'un premier signal de prédiction représentatif d'une prédiction d'un desdits premiers signaux de composantes spectrales, ladite prédiction étant basée sur un ou plusieurs desdits premiers signaux de composantes spectrales;

(c) comparaison du premier signal de prédiction avec ledit un desdits premiers signaux de composantes spectrales afin de générer un premier signal d'erreur de prédiction;

(d) exécution d'une décomposition spectrale d'un deuxième signal de canal audio en une pluralité de deuxième signaux de composantes spectrales;

(e) exécution d'un codage stéréo combiné dudit un desdits premiers signaux de composantes spectrales et d'un desdits deuxième signaux de composantes spectrales afin de générer un signal de composantes spectrales codé combiné, ledit codage étant basé sur le premier signal d'erreur de prédiction; et

(f) génération du signal codé sur la base du signal de composantes spectrales codé combiné.

2. Procédé selon la revendication 1, comprenant en outre les étapes suivantes:

(g) génération d'un deuxième signal de prédiction représentatif d'une prédiction dudit un desdits deuxième signaux de composantes spectrales, ladite prédiction étant basée sur un ou plusieurs autre(s) desdits deuxième signaux de composantes spectrales; et

(h) comparaison du deuxième signal de prédiction avec ledit un desdits deuxième signaux de composantes spectrales afin de générer un deuxième signal d'erreur de prédiction;

et dans lequel l'étape d'exécution du codage stéréo combiné dudit un desdits premiers signaux de composantes spectrales et dudit un desdits deuxième signaux de composantes spectrales est en outre basée sur ledit deuxième signal d'erreur de prédiction.

3. Procédé selon la revendication 1, dans lequel l'éta-

- pe d'exécution du codage stéréo combiné dudit un desdits premiers signaux de composantes spectrales et dudit un desdits deuxièmes signaux de composantes spectrales comprend l'exécution d'un codage stéréo d'intensité dudit un desdits premiers signaux de composantes spectrales et dudit un desdits deuxièmes signaux de composantes spectrales.
4. Procédé selon la revendication 1, dans lequel l'étape de génération du signal codé sur la base du signal de composantes spectrales codé combiné comprend la quantification du signal de composantes spectrales codé combiné.
5. Procédé selon la revendication 4, dans lequel ladite quantification du signal de composantes spectrales codé combiné est basée sur un modèle perceptif.
6. Procédé selon l'une quelconque des revendications précédentes, comprenant en outre l'étape de stockage dudit signal codé sur un support de stockage.
7. Procédé selon la revendication 6, dans lequel le support de stockage comprend un disque compact.
8. Procédé selon la revendication 6, dans lequel le support de stockage comprend une bande audio numérique.
9. Procédé selon la revendication 6, dans lequel le support de stockage comprend une mémoire à semiconducteur.
10. Procédé de découpage d'un signal codé afin de générer un signal audio multicanaux reconstruit, le signal codé comprenant un codage stéréo combiné d'un signal audio multicanaux original, le procédé comprenant les étapes suivantes:
- (a) exécution d'un décodage stéréo combiné du signal codé afin de générer une pluralité de signaux de canal décodés, chaque signal de canal décodé comprenant une pluralité de signaux d'erreur de prédiction de composantes spectrales décodés;
 - (b) génération d'un premier signal de composantes spectrales basé sur un ou plusieurs desdits signaux d'erreur de prédiction de composantes spectrales compris dans un premier desdits signaux de canal décodés;
 - (c) génération d'un premier signal de prédiction représentatif d'une prédiction d'un deuxième signal de composantes spectrales, ladite prédiction étant basée sur ledit premier signal de composantes spectrales;
 - (d) génération du deuxième signal de composantes spectrales sur la base du premier signal de prédiction et d'un ou de plusieurs desdits signaux d'erreur de prédiction de composantes spectrales compris dans le premier desdits signaux de canal décodés; et
 - (e) génération d'un premier canal du signal audio multicanaux reconstruit sur la base des premier et deuxième signaux de composantes spectrales.
11. Procédé selon la revendication 10, comprenant les étapes suivantes:
- (f) génération d'un troisième signal de composantes spectrales sur la base d'un ou de plusieurs desdits signaux d'erreur de prédiction de composantes spectrales compris dans un deuxième desdits signaux de canal décodés;
 - (g) génération d'un deuxième signal de prédiction représentatif d'une prédiction d'un quatrième signal de composantes spectrales, ladite prédiction étant basée sur ledit troisième signal de composantes spectrales;
 - (h) génération du quatrième signal de composantes spectrales sur la base du deuxième signal de prédiction et d'un ou de plusieurs desdits signaux d'erreur de prédiction de composantes spectrales compris dans le deuxième desdits signaux de canal décodés; et
 - (i) génération d'un deuxième canal du signal audio multicanaux reconstruit sur la base des troisième et quatrième signaux de composantes spectrales.
12. Procédé selon la revendication 10, dans lequel l'étape d'exécution du décodage stéréo combiné du signal codé comprend l'exécution d'un décodage stéréo d'intensité du signal codé.
13. Codeur pour exécuter un codage stéréo combiné d'un signal audio multicanaux afin de générer un signal codé, le codeur comprenant:
- (a) un premier banc de filtres (12L, 12R), qui exécute une décomposition spectrale d'un premier signal de canal audio en une pluralité de premiers signaux de composantes spectrales;
 - (b) un premier filtre de prédiction (16L, 16R, 47), qui génère un premier signal de prédiction représentatif d'une prédiction d'un desdits premiers signaux de composantes spectrales, ledit filtre de prédiction étant sensible à un ou à plusieurs autre(s) desdits premiers signaux de composantes spectrales;
 - (c) un premier comparateur (48), qui compare le premier signal de prédiction avec ledit un desdits premiers signaux de composantes spectrales afin de générer un premier signal d'erreur de prédiction;

- (d) un deuxième banc de filtres (12L, 12R), qui exécute une décomposition spectrale d'un deuxième signal de canal audio en une pluralité de deuxièmes signaux de composantes spectrales; 5
- (e) un codeur stéréo combiné (13), qui exécute un codage stéréo combiné dudit un desdits premiers signaux de composantes spectrales et d'un desdits deuxièmes signaux de composantes spectrales afin de générer un signal de composantes spectrales codé combiné, ledit codage étant basé sur le premier signal d'erreur de prédiction; et 10
- (f) un codeur (14), qui génère le signal codé sur la base du signal de composantes spectrales codé combiné. 15
- 14.** Codeur selon la revendication 13, comprenant en outre: 20
- (g) un deuxième filtre de prédiction (16L, 16R, 47), qui génère un deuxième signal de prédiction représentatif d'une prédiction dudit un desdits deuxièmes signaux de composantes spectrales, ladite prédiction étant basée sur un ou plusieurs autre(s) desdits deuxièmes signaux de composantes spectrales; et 25
- (h) un deuxième comparateur (48), qui compare le deuxième signal de prédiction avec ledit un desdits deuxièmes signaux de composantes spectrales afin de générer un deuxième signal de prédiction d'erreur; 30
- et dans lequel le codeur stéréo combiné exécute un codage stéréo combiné basé en outre sur ledit deuxième signal d'erreur de prédiction. 35
- 15.** Codeur selon la revendication 13, dans lequel le codeur stéréo combiné comprend un codeur stéréo d'intensité, qui exécute un codage stéréo d'intensité dudit un desdits premiers signaux de composantes spectrales et dudit un desdits deuxièmes signaux de composantes spectrales. 40
- 16.** Codeur selon la revendication 13, dans lequel le codeur qui génère le signal codé sur la base du signal de composantes spectrales codé combiné comprend un quantificateur, qui quantifie le signal de composantes spectrales codé combiné. 45
- 17.** Codeur selon la revendication 16, dans lequel le quantificateur est basé sur un modèle perceptif. 50
- 18.** Décodeur pour le décodage d'un signal codé afin de générer un signal audio multicanaux reconstruit, le signal codé comprenant un codage stéréo combiné d'un signal audio multicanaux original, le procédé comprenant : 55
- (a) un décodeur stéréo combiné (23), qui exécute un décodage stéréo combiné du signal codé afin de générer une pluralité de signaux de canal décodés, chaque signal de canal décodé comprenant une pluralité de signaux d'erreur de prédiction de composantes spectrales décodés;
- (b) des moyens pour générer un premier signal de composantes spectrales basé sur un ou plusieurs desdits signaux d'erreur de prédiction de composantes spectrales dans un premier desdits signaux de canal décodés;
- (c) un premier filtre de prédiction (26L, 26R), qui génère un premier signal de prédiction représentatif d'une prédiction d'un deuxième signal de composantes spectrales, ladite prédiction étant basée sur ledit premier signal de composantes spectrales;
- (d) des moyens pour générer le deuxième signal de composantes spectrales sur la base du premier signal de prédiction et d'un ou de plusieurs desdits signaux d'erreur de prédiction de composantes spectrales compris dans le premier desdits signaux de canal décodés; et
- (e) un premier banc de filtres (24L, 24R), qui génère un premier canal du signal audio multicanaux reconstruit sur la base des premier et deuxième signaux de composantes spectrales.
- 19.** Décodeur selon la revendication 18, comprenant en outre:
- (f) des moyens pour générer un troisième signal de composantes spectrales sur la base d'un ou de plusieurs signaux de prédiction d'erreur de composantes spectrales compris dans un deuxième desdits signaux de canal décodés;
- (g) un deuxième filtre de prédiction, qui génère un deuxième signal de prédiction représentatif d'une prédiction d'un quatrième signal de composantes spectrales, ladite prédiction étant basée sur ledit troisième signal de composantes spectrales;
- (h) des moyens pour générer le quatrième signal de composantes spectrales sur la base du deuxième signal de prédiction et d'un ou de plusieurs des signaux d'erreur de prédiction de composantes spectrales compris dans ledit deuxième desdits signaux de canal décodés; et
- (i) un deuxième banc de filtres, qui génère un deuxième canal du signal audio à canaux multiples reconstruit sur la base des troisième et quatrième signaux de composantes spectrales.
- 20.** Décodeur selon la revendication 18, dans lequel le décodeur stéréo combiné comprend un décodeur

stéréo d'intensité, qui exécute un décodage stéréo d'intensité du signal codé.

5

10

15

20

25

30

35

40

45

50

55

FIG. 1
(PRIOR ART)

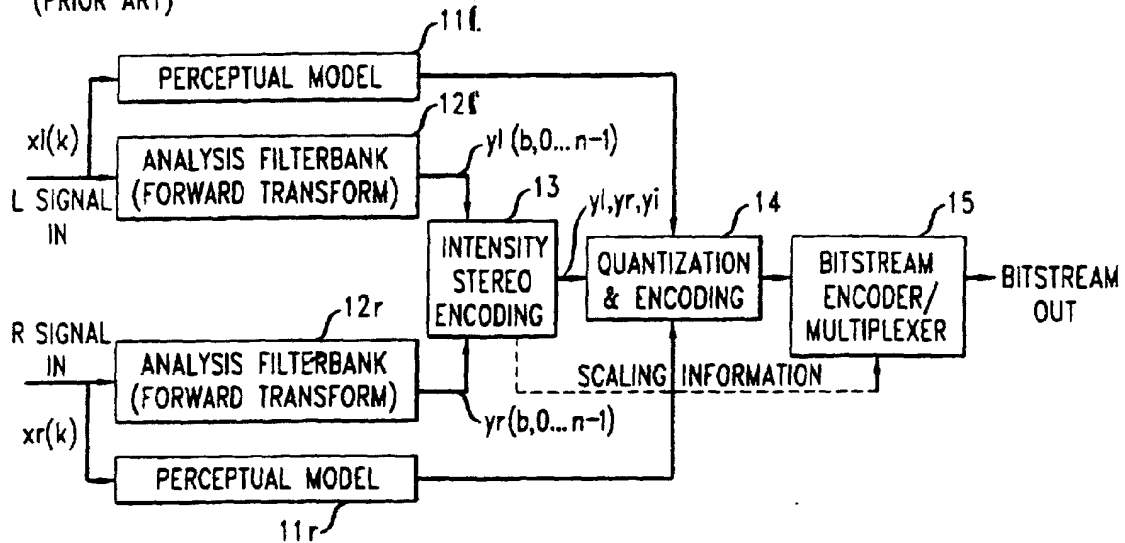


FIG. 2

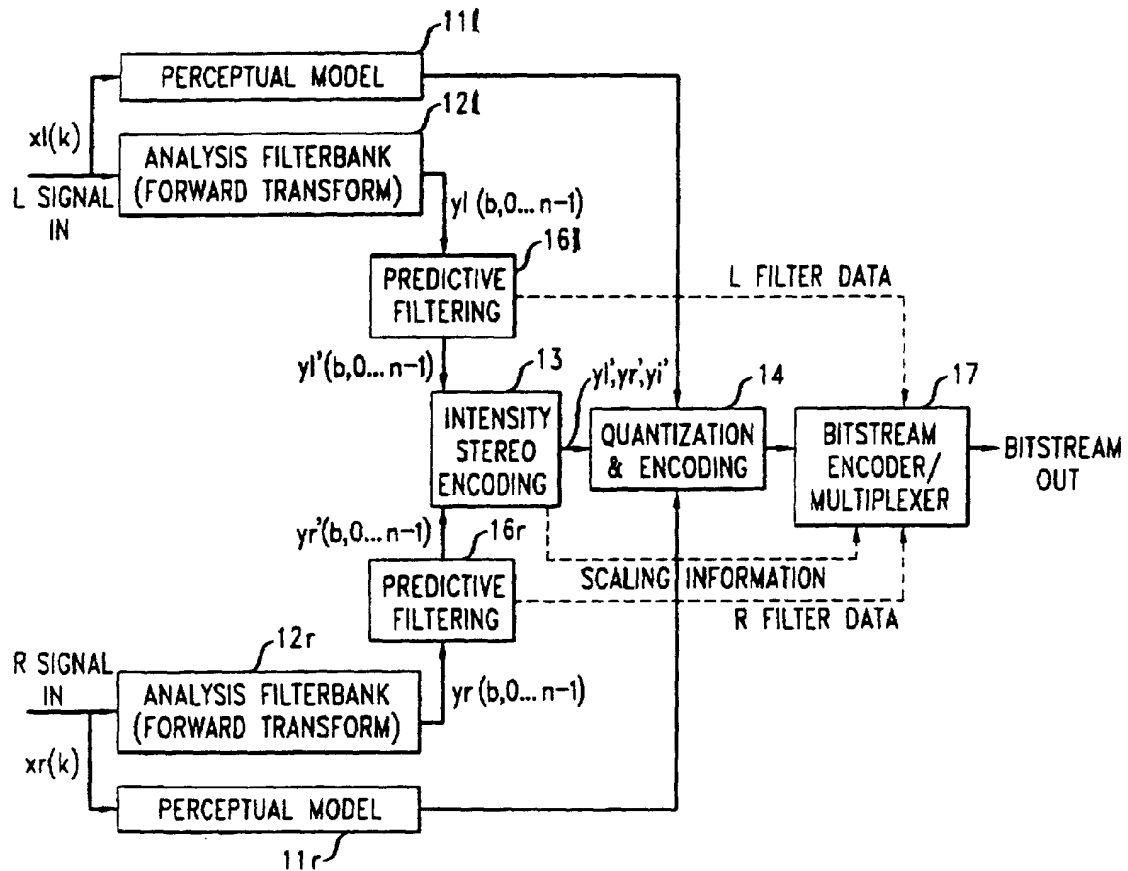


FIG. 3

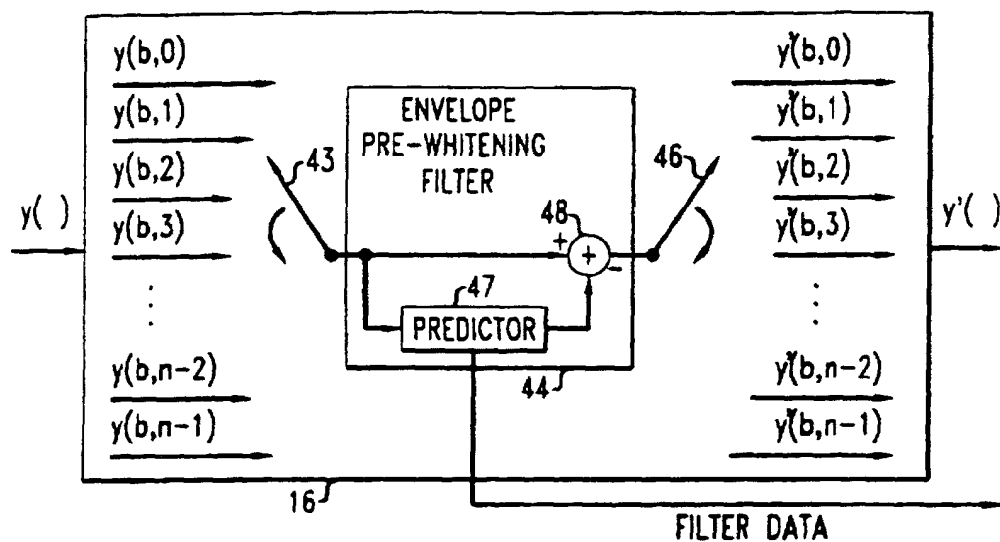
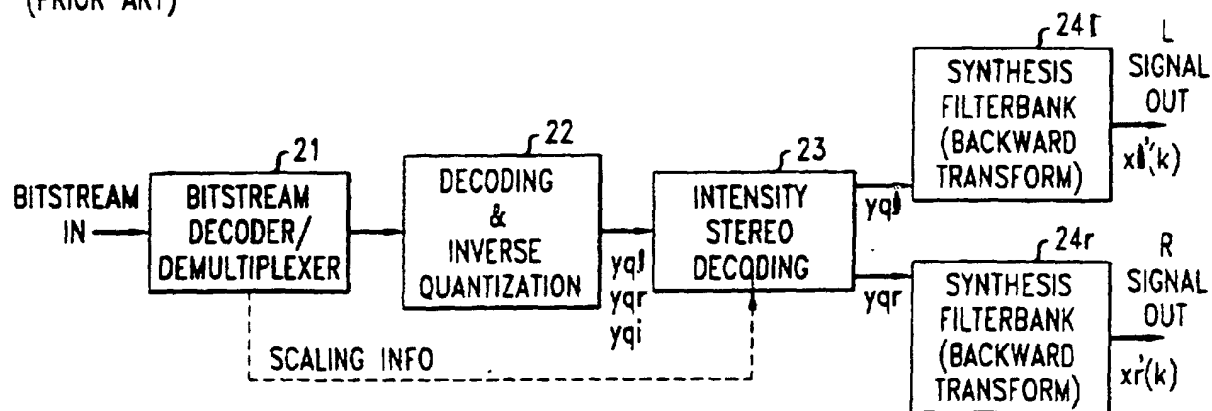
FIG. 4
(PRIOR ART)

FIG. 5

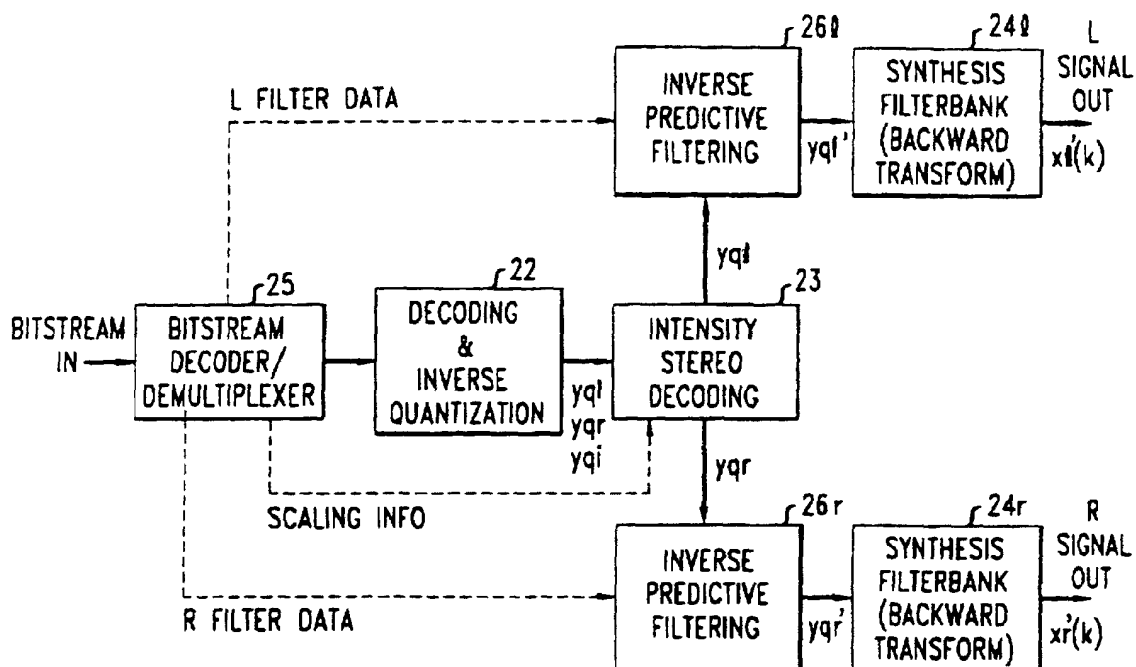


FIG. 6

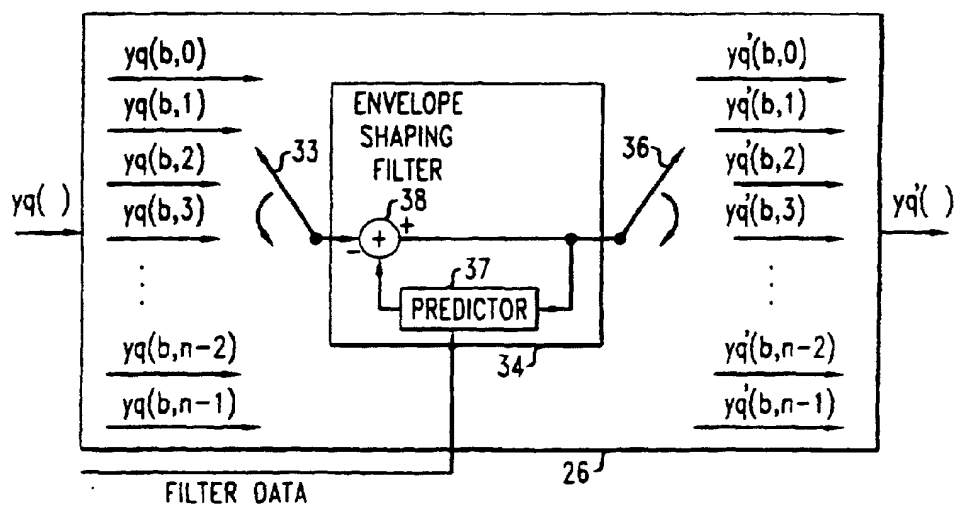


FIG. 7

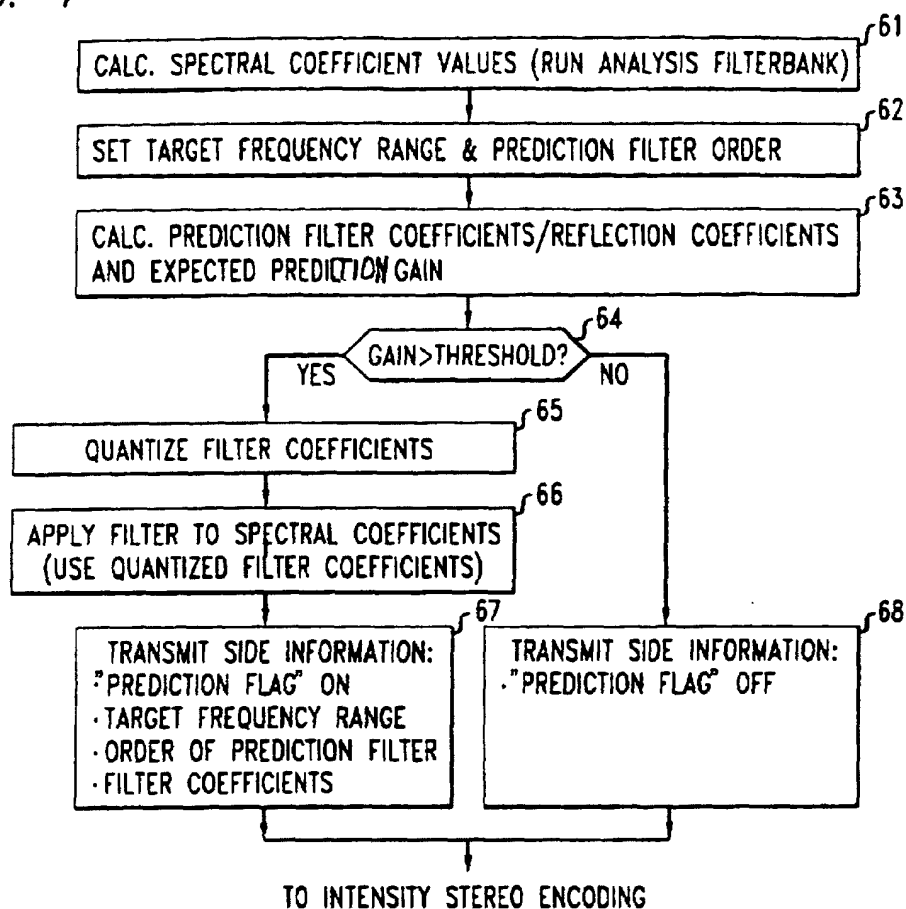


FIG. 8

